Forensic Analysis in a Digital World
Gary L. Palmer, INFOSEC Scientist, The MITRE Corporation

Perception

World cultures have formed ever increasing dependencies on digital systems and networks. As such, digital technology is becoming commonplace and in some cases necessary in many people's normal day-to-day activities. It stands to reason that, much like other cultural changes that have moved in to modify our lives, digital technology will increasingly be used by anti-social and nefarious types as well as normal citizens in our expanding digital world.

The basic stages involved in adapting society to the wrongful use of many technologies are as follows. First, is a realization by consumers of the technology that it can and is being used for unauthorized and possibly unlawful purposes. Growing concern as incidents rise and become more serious then follows realization. The volume of misuse and percentage of unlawful activity will eventually cause authorities to recognize that they need some level of expertise to help identify, understand and possibly thwart any future wrongdoing. This stage is preceded by cultivating certified expertise supported by an ever-deeper understanding of the problem, its symptoms and the motivations of those involved in wrongful use. How thoroughly these stages are implemented is a function of several related factors. Two of those factors are, perhaps, most important. First, is the complexity involved in the technology. True subject matter experts are required to understand the associated technology completely as a prerequisite to stating opinions or conclusions about evidence. Second, sufficient conclusive research must stand behind techniques and methods (including tools) employed to analyze and examine exhibits that could become evidence or proof. Up to the present, actions to address both of these related issues have been closely aligned with the formation and evolution of most "forensic" disciplines.

Due, in large part, to our focus on entertainment in western society such as “Quincy ME” or Discovery Channel’s “The New Detectives” or “CSI: Crime Scene Investigators,” the word “forensic” conjures up specific images. Most everyone, even those in the scientific community, have a pre-conceived notion of the discipline called forensic analysis, and it usually involves visions of autopsies, DNA analysis, and men and women in white coats in a cold, sterile lab. The domain of analysis performed at a "brick and mortar" crime lab, by highly trained (and sworn in most cases) practitioners on tangible, physical items found on, in or around a body at the scene of a crime involving death or terrible injury and solely in support of Law Enforcement and the courts.

With that in mind, where does the emerging discipline called Computer Forensics or the even less understood area named Network Forensics fit? Aside from the fact that in many cases a physical computer or hard drive has been seized and shipped to a "brick and mortar" lab, one has to wonder where the connection is between forensic analysis and the digital systems and digital networks so prevalent today. The latter seems to call for near real-time techniques applied to active systems and networks that strive to be predictive or anticipatory. The former appears to reside in an ex post facto world where exhibits are seized, sent to a cloistered facility, and slowly, meticulously analyzed, a totally reactive process. In either case very little of the potential evidence is tangible, or physical. Rather, it is highly interpreted and subject to complex transformations that act on the raw data to place it in a form that can be scrutinized or analyzed.
Computers and other digital components are very complex systems. The product of decades of engineering and manufacturing improvements supported by centuries of scientific study and academic research, computers are truly marvelous devices with great positive potential. This potential makes it certain that digital systems will be used in positive and negative ways. So, as our culture's authorities step through the stages to form the expertise they need to recognize and stop wrongdoing, two important issues must be addressed. Experts must understand the complexity involved in digital technology and that must be aided by serious academic research as a prelude to effective tools and techniques applied in the analysis of digital systems. This paper contends that, up to the present, the evolution of digital forensics has taken a different path. Primarily due to urgent need recognized by analysts and examiners, the field has grown somewhat in reverse of convention. Tools and techniques came before research and expert cultivation. My contention is that for our field to become a true forensic discipline this trend must be reversed. Defense attorneys are beginning to question long established precedent, and courts are increasingly calling for scientific and technical evidence to be judged by rigorous standards (Pollack, 2002). Part of the solution lies in the realization that although forensic analysis is commonplace in support of criminal cases, it is not necessarily under the sole purview of Law Enforcement. Rather, it should be viewed as a rigorous scientific specialty whose purpose is to provide information “suitable” for the courts or public forum. This definition allows for a broader application that will be explored in some detail over the next few pages.

Historical View

Given their long history and current success in providing factual, testimonial evidence for the courts, it is prudent to begin by referencing more traditional methods in forensic science. Review of other forensic analysis methods will help in understanding how we can apply similar techniques when dealing with information systems. Traditional methods include:

- chromatography, spectroscopy, hair and fiber analysis, serology (DNA);
- pathology, anthropology, odontology, toxicology; structural engineering, and questioned documents;
- behavioral patterns revealed by tests such as polygraphs and psychological battery exams.

Most of these forensic disciplines began to flourish alongside the evolving science of criminalistics which, in the United States, emerged during the 1920’s. Advances in medicine, chemistry and microscopy prepared the way for the adoption of scientific analysis, rather than pure observation and intuition as the cornerstone of criminal investigation. The result of this advancement was, of course, to replace supposition with reality (or fact) and present testimonial evidence to the trier-of-fact (judge or jury) in criminal or civil proceedings.

The vast majority of analytical methods employed by traditional forensic sciences grew out of university laboratories. In fact, before 1929 no official crime laboratory even existed in the United States. Instead, police departments interested in using scientific analysis in the solving of crimes would solicit the assistance of prominent university professors to help them collect and examine potential evidence (Eckert, 1997). Over time, as more and more federal, state and local jurisdictions realized the importance and necessity of scientific investigation, professionals with particular interest in the forensic aspects of analysis transitioned their practices to newly established laboratories that focused on forensic analysis in support of the courts. This trend remains true to this day, although, as stated previously, forensic analysis of computer systems has taken a different evolutionary path.
The gradual paradigm shift, from intuition or supposition to fact derived from analysis, took
hold in the early twentieth century for a number of reasons. The sciences, both hard (physics)
and soft (biology), were advancing rapidly and many of their discoveries were being exposed to a
larger percentage of the common population. Perhaps more important was the fact that surface
observation alone had been proven time and time again to lead to suspect conclusions. Over
time, conclusions presented as scientific evidence in the courts became subject to more rigorous
scrutiny. The court system realized that testimony proffered as scientific and conclusive was, for
the most part, beyond their complete understanding. In addition, the courts also understood that
these analytical methods were not irrefutable. They were derived by experimentation that
contained (or should contain) measures of error and other indices that help describe the veracity
of statistics and narrative results. This leads to standards and rules of admissibility as well as the
expert testimony that must accompany scientifically derived testimonial evidence (Eckert, 1997).

Mostly in criminal proceedings, the courts, and public opinion, have come to rely heavily on
certain evidence derived by the scientific method. Perhaps the most commonly stated, but least
understood, is DNA profiling. This relatively new method is performed for the courts as a
technique used by forensic serologists. It is relied upon because of its purported ability to
discriminate down to the individual thus replacing other, older, methods like blood typing as a
primary evidentiary mechanism. Looking a little deeper DNA analysis, though certainly more
reliable than typing alone, is not a panacea. The general assumption is that presenting DNA
evidence in court is irrefutable and can therefore not be contested. This supposition is founded
upon studies based on population genetics where false positive rates are exceptionally small (i.e.
one in billions) or stated another way, a reliance on the probability that the DNA analysis will
correctly determine that a defendant was the source of evidence found at the crime scene.
However, when statistics that take laboratory practice and data collection factors into
consideration are gathered, false positive detection rates range from one per hundred to one per
thousand (Koehler, 1995). This begins to approach the false positive rates for blood typing. This
view of DNA evidence seems much more applicable to courts since they are serviced by
laboratories like those studied (Koehler). It renders the studies based on population genetics
potentially irrelevant since so much error can be interjected by incorrect collection and handling
of the DNA source material.

One lesson to be learned is that there is error in every analysis method. There is no doubt that
the scientific community has agreed that DNA profiling is very accurate. The question that
remains relates to the reliability of any particular test. This same question is a pivotal issue for
current and future practitioners in computer forensic analysis. Are bits dropped on the floor
during an imaging operation? If so, is there a measurable frequency of that occurrence and is it
statistically significant? Are collection tools missing or not reporting exculpatory data? If yes,
when, under what circumstances or conditions? Is the algorithm for verifying graphic format
missing all ART files? Is it checking file types rather that using a "magic" file to read and verify
file headers and trailers? These are but a tiny sample of questions that should be asked and
addressed to help understand these complex digital systems and verify the tools and techniques
we will use. The point to be made is that error rates in analysis are a fact. They should not be feared, but they must be measured. That is one of the reasons experiments are performed. Until
very recently, the scientific community has been conspicuously absent in the development of
standards, processes and protocols related to forensic analysis of digital components (CFTT, 2002). This has lead to the court’s reliance on precedent rather than statistical significance and repeatability when ruling on admissibility of evidence derived from digital sources. As judges, juries, defense attorneys and asset managers become better schooled in digital technology and understand its complexity more completely, it is likely that we will see the call for a more rigorous approach to analysis. Once this begins decision makers will ask more compelling questions and expect more detailed, scientifically proven explanations from those providing testimony or persuasive argument. This new view of evidence coupled with increasingly dynamic, networked environments will force a paradigm shift. This shift will slowly change law enforcement’s use of technology and allow for a wider use of forensic techniques in business, industry, government, and the military. In fact, we have already seen the start of this trend in recent judicial opinions that site the need for more rigorous science as criteria for admissibility (Pollack).

Although viewed, initially, as troublesome, the benefit of adding rigor to the collection analysis and presentation of scientific evidence will result in much higher confidence levels associated with the information presented to all decision-makers including judges and juries. For the digital forensic analyst working in near-real-time environments, it will allow for quicker responses based upon more reliable evidence derived from proven technology grounded in accepted standards. The goal is to produce reliable information that serves to maintain continuity of operations, while at the same time possessing characteristics that make it suitable for presentation in the courts.

Solution Path

The forensic analysis of computer systems, whether it be for the trier-of-fact in the courts or decision makers in business or military operation, has the same goal; persuasion based on factual evidence. The information must be sufficient to help commit a judge and jury to a verdict, or it must help allow a decision maker to change resource allocations or operational goals (and accept residual risk). In the courts, for information to have the opportunity to persuade it must first be admitted. In business and the military decision makers must have confidence in the messenger and the mechanism. At the core they are essentially identical, just called by different names. In general, the information must be:

- **Relevant and/or Material**: will this information assist the decision maker in his task?
- **Credible and/or Competent**: is the information believable, trustworthy, true and, if so, by what measure?

Some subset of these characteristics are applied to all information offered to persuade. In the fledgling science of forensic analysis of information systems this is becoming more evident. Whereas the traditional forensic sciences have long established histories and defined laboratory protocols for tests, professional advocacy groups, and university support, computer forensic analysis has only recently come on the scene in response to undefined, illegal use of readily available technology. In a sense, we are where the other disciplines were in the early part of the twentieth century: an evolving scientific discipline, becoming more familiar to the general populace, and searching for measures of accuracy and reliability so as to increase confidence and credibility.
Measures of reliability and accuracy for the techniques and methods used in analysis goes to the level of confidence expected in the evidence and accompanying testimony. Information derived from computer forensic analysis has yet to be contested to any great extent by defense lawyers in judicial proceedings or analysts in investigations of computer misuse. Most techniques used today are assumed plausible if not incontestable because they are developed by reputable companies, used by experts or practitioners in the field, and have been used in courts, or to otherwise persuade authority, before. The techniques themselves and the conclusions they lead to have yet to be tested for reliability in controlled environments under experimental protocols. Strict interpretations of the rules of evidence and court precedence imply that this will soon be necessary if digital evidence is taken to task (Foster 1997, FRE 2000). The complexity evident in digital systems will make this a very difficult road, one that our community of researchers and practitioners should at least be getting ready to travel on soon.

Parting Thoughts

Some of the methods employed by the traditional forensic sciences have much to teach those interested in the new field of computer forensic analysis. Using DNA as an example once again, one can see the cumulative effect of discovery through the years. Johann Meischer's analysis of old bandages in the Crimean War lead to his discovery of what he called "nuclein" in 1869. Watson and Crick defined the structure of DNA in 1953, which gave researchers a blueprint. Gilbert and Sanger described how to sequence DNA in 1977, which allowed researchers to analyze small parts of the structure. Alex Jeffery found the uniquely human part of the strand in 1985, which made unique comparisons possible. These four distinct events along with hundreds occurring over 116 years represent how scientific discovery works. If we expect computer forensics to join serology in the ranks of proven forensic disciplines, then we should expect similar processes to be at work with strong interaction among academic research, field practitioners, and legal experts.

In the courts, admission and presentation of scientific evidence is guided by established judicial rule and legal precedence. It stands to reason that evidence analyzed from computer systems will, in the near future, be called upon to meet the same exacting standard. So even though we can do binary analysis with hashing algorithms to analyze the very fibers of the computer system itself, it will be the accuracy and reliability of the hash employed that may be called into question. We can claim to use proven correct tools to do an ‘autopsy’ on a computer system after it has been compromised. The questions will be “Define proven correct?” and “Using what standard?”

These issues only get more complicated as we move from a single host at a physical crime scene to the “virtual crime scene,” which consists of networked systems, and devices located anywhere in our infosphere. No matter what the environment, the need for admissible, conclusive evidence will be required and must be collected from all sources available. This includes the subject or compromised host itself, as well as distant firewalls, routers, smart hubs, application gateways, wireless devices, cellular components, deployed agents, and intrusion detection systems. In the near future, the collection, fusion and correlation of data from all these sources and more will be vital to investigations, both civil and criminal. It will be increasingly important that evidence and the methods and techniques used to uncover it are accurate, reliable, and accepted as standard in our field. Coupled with certified expertise, the incorporation of the
scientific method is the key to providing forensic evidence or suitable information meant to persuade, whether it is for courts of law, military operations, e-commerce or homeland defense.

References


© 2002 International Journal of Digital Evidence

About the Author

Gary Palmer (palmerg@mitre.org) works as a Senior INFOSEC Scientist for the MITRE corporation out of Orlando, Florida where he supports the Air Force Research Laboratory (ARFL) programs related to Digital Forensic Analysis, Intrusion Detection and Defensive Information Warfare. Over the last 22 years Gary has worked in the areas of large systems development, integration, database design, software development and information assurance. He is currently lead coordinator for AFRL’s Digital Forensic Research Workshop and R&D lead for the CyberForensic Science and Technology Center.